

WE CLAIM:

1. A method of employing oligonucleotide probes to obtain information on a target nucleic acid analyte containing a target sequence segment, the method comprising:

5 contacting the analyte, under hybridizing conditions, with at least two oligonucleotide probes, each oligonucleotide probe comprising a sequence segment complementary, or complementary except at a position corresponding to a probed position of the target sequence,

10 wherein a nucleotide at the position of each oligonucleotide probe corresponding to the probed position is capable base pairing with a set of two or more nucleotides, each set is unique but includes one nucleotide common to all the sets, and one nucleotide that may be present in the target sequence segment is not represented in any set,

15 further wherein hybridization of each oligonucleotide probe to the target sequence segment under the hybridizing conditions occurs only if no mismatch exists at the probed position, such that depending upon the identity of the nucleotide at the probed position of the target sequence segment, all, some or none of the oligonucleotide probes hybridize to the target nucleic acid sequence.

20 2. The method of claim 1 wherein four nucleotides may be present in the target sequence segment and each oligonucleotide probe comprises, at the position corresponding to the probed position a nucleotide base pairing with two nucleotides.

25 3. The method of claim 1 wherein more than four nucleotides may be present in the target sequence segment, each oligonucleotide probe comprises, at the position corresponding to the probed position a nucleotide base pairing with more than two nucleotides, and more than two oligonucleotide probes are employed.

30 4. The method of claim 3 wherein five nucleotides may be present in the target sequence segment, three oligonucleotide probes are employed, each oligonucleotide probe comprises, at the position corresponding to the variable position a nucleotide base pairing with two or three nucleotides, and each set has at least one nucleotide not common to all the sets in common with another set.

5. The method of claim 1 wherein a null hybridizing sequence comprising a nucleic acid sequence, complementary to the nucleic acid sequence of interest, having the nucleotide represented in neither the first or second set of two or more nucleotides at the variable position is employed.

5

6. The method of claim 1 employed as a sequencing method.

7. The method of claim 6 wherein the sequencing method is by analysis of hybridization data obtained from an array of oligonucleotide probes.

10

8. The method of claim 7 wherein the array comprises arrayed individual beads or particles, each bead or particle having a surface to which is attached a plurality of oligonucleotide probes of identical sequence.

15

9. The method of claim 7 wherein the array comprises a substrate having a surface, the surface having a plurality of discrete surface sites, each site having attached a plurality of oligonucleotide probes of identical sequence.

20

10. The method of claim 7 wherein detection of a target sequence segment hybridizing to an oligonucleotide probe is by detection of a discrete label moiety linked to the target sequence segment.

25

11. The method of claim 10 wherein the discrete label moiety linked to the target sequence segment comprises a nucleic acid sequence.

12. The method of claim 10 wherein the discrete label moiety linked to the target sequence segment comprises a luminescent moiety.

30

13. The method of claim 12 wherein the luminescent moiety is a chemiluminescent or fluorescent moiety.

14. The method of claim 9 wherein detection of a target sequence segment hybridizing to an oligonucleotide probe is by detection of an label intrinsic to the target sequence segment.

5 15. The method of claim 14 wherein the label intrinsic to the target sequence segment is ³²P.

16. The method of claim 7 wherein detection of a target sequence segment hybridizing to an oligonucleotide probe is by detection of the heat of hybridization.

10

17. The method of claim 6 wherein the sequencing method is by detection of labels that attach to by hybridization to the target sequence segment.

15 18. The method of claim 1 wherein hybridized target nucleic acids are amplified by a polymerase enzyme that requires a hybridized complex for polymerizing the formation of nucleic acid sequence.

19. The method of claim 18 wherein hybridized nucleic acids are amplified by polymerase chain reaction.

20

20. The method of claim 19 wherein hybridized nucleic acids are amplified by an RNA replicase enzyme.

21. The method of claim 19 for genetic analysis.

25

22. The method of claim 21 employed for allelic analysis.

23. The method of claim 21 wherein genomic DNA is analyzed.

30

24. The method of claim 21 wherein genomic cDNA is analyzed.

25. The method of claim 2 wherein the four nucleotides are A, T, C and G.

26. The method of claim 4 wherein the five nucleotides are A, T, C, G and I (Inosine).

27. The method of 6 wherein the sequencing method is by analysis of
5 hybridization data obtained from an array of target nucleic acid analyte sequences attached to a substrate surface.

28. The method of 14 wherein the array comprises arrayed individual beads or particles, each bead or particle having a surface, the surface having attached a plurality of
10 target nucleic acid analyte sequences having an identical sequence.

29. The method of 14 wherein the array comprises arrayed discrete sites on a substrate surface of an integrated substrate, each site having a surface, the surface having attached a plurality of target nucleic acid analyte sequences having an identical sequence.
15

30. The method of claim 14 wherein each oligonucleotide probe sequence additionally comprises a linker moiety and a label moiety.

31. The method of claim 15 wherein the linker moiety comprises a common
20 nucleic acid sequence and the label moiety comprises a signature nucleic acid sequence that identifies the target sequence segment.

32. The method of claim 16 wherein the common nucleic acid sequence is double stranded.
25

33. The method of claim 17 wherein decoder labels comprising a nucleic acid sequence complementary to the signature sequence and a second label moiety are employed to image the array.

30 34. The method of claim 18 wherein the second label moiety comprises a luminescent moiety.

35. The method of claim 34 wherein the luminescent moiety is a fluorescent or chemiluminescent moiety.

36. The method of claim 14 wherein the substrate surface is functionalized with
5 a surface modification to enhance hybridization.

37. The method of claim 36 wherein the enhancement is increasing stringency or kinetics of hybridization.

10 38. The method of claim 14 wherein the electric potential at the substrate surface is electronically controlled to enhance hybridization.

39. The method of claim 29 wherein the integrated substrate comprises a semiconductor chip comprising electronic circuitry, wherein the electric potential at the
15 individual array sites of the substrate surface is independently electronically controlled to enhance of hybridization.

40. A method for analysis of a plurality of target nucleic acid sequences, using label nucleotide sequences, by base pairing complementarity under hybridizing
20 conditions that reduce the number of label nucleotide sequences required to uniquely label each of the target nucleotide sequences, the method comprising:

a) contacting each target nucleic acid sequence with a collection of label nucleotide sequences each comprising a label moiety and an anti-target sequence segment complementary, or complementary except at one variable position, to a target
25 sequence segment of the target nucleic acid sequence,

wherein the position of each anti-target sequence segment corresponding to the variable position comprises a nucleotide base pairing with a set of two or more nucleotides present in the plurality of nucleic acid sequences, each set is different from each other set, each set includes at least one nucleotide in common with each other set,
30 and a nucleotide present in the plurality of nucleic acid sequences is not represented in any set, and hybridization of the anti-target sequence segment of each label nucleotide sequence to a target sequence segment under the conditions occurs only if

complementarity exists between the nucleotide at the variable position and the nucleotide at the corresponding position of each label nucleotide; and

b) detecting which of the label nucleotide sequences hybridize to each target nucleic acid sequence,

5 wherein depending upon the identity of the nucleotide at the variable position of the target sequence segment, some, all or none of the label nucleotide sequences employed hybridize to each target nucleotide sequence.

41. The method of claim 40 wherein four nucleotides are present in the sequence
10 of interest and, two label nucleotide sequences are employed for each target sequence segment, and each anti-target sequence segment comprises, at the position corresponding to the variable position of the target sequence segment, a nucleotide base pairing with two nucleotides.

15 42. The method of claim 40 wherein more than four nucleotides are present in the sequence of interest, each anti-target sequence segment comprises, at the position corresponding to the variable position of the anti-target sequence, a nucleotide base pairing with more than two nucleotides, and the collection of label nucleotide sequences includes more than two label nucleotide sequences.

20 43. The method of claim 42 wherein five nucleotides may be present in the target sequence segment, three label nucleotide sequences are employed, each label nucleotide sequence comprises, at the position corresponding to the variable position of the anti-target sequence, a nucleotide base pairing with two or three nucleotides and each set has
25 at least one nucleotide not common to all the sets in common with another set.

44. The method of claim 40 wherein a null label nucleotide sequence is employed, the null label nucleotide sequence comprising a nucleic acid sequence complementary to a target segment of a target nucleic acid sequence, having the
30 nucleotide not represented in any set of two or more nucleotides at the variable position.

45. The method of claim 40 employed as a sequencing method.

46. The method of claim 45 wherein the sequencing method is by analysis of hybridization data obtained from an array of label nucleotide sequences.

47. The method of claim 46 wherein the array comprises arrayed individual
5 beads or particles, each bead or particle having a surface to which is attached a plurality of label nucleotide sequences having an identical sequence.

48. The method of claim 46 wherein the array comprises a substrate having a surface, the surface having a plurality of discrete surface sites, each site having attached
10 a plurality of label nucleotide sequences having an identical sequence.

49. The method of claim 46 wherein detection of a target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of a discrete analyte label moiety linked to the target nucleic acid sequence.
15

50. The method of 49 wherein identification of the label nucleotide sequence to which a target nucleic acid sequence hybridizes is by identification of the label moiety of the label nucleotide sequence.

51. The method of claim 49 wherein the analyte label moiety linked to the target nucleic acid sequence comprises a nucleic acid sequence.
20

52. The method of claim 49 wherein the discrete analyte label moiety linked to the target nucleic acid sequence comprises a luminescent moiety.
25

53. The method of claim 52 wherein the luminescent moiety is a chemiluminescent or fluorescent moiety.

54. The method of claim 53 wherein the discrete analyte label moiety of the target nucleic acid sequence and the label moiety of the label nucleotide sequence are
30 both fluorescent and capable of fluorescence resonance energy transfer, and detection of hybridization of label nucleotide sequence and target nucleic acid sequence is by detecting fluorescence resonance energy transfer.

55. The method of claim 47 wherein detection of a target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of an analyte label intrinsic to the target nucleic acid sequence of interest.

5

56. The method of claim 40 wherein a discrete analyte label moiety is linked to each target nucleic acid sequence, identification of the label nucleotide sequence hybridizing to a target nucleic acid sequence is by detection of the label moiety, and identification of the target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of the discrete analyte label moiety linked to the target nucleic acid sequence.

10

57. The method of claim 55 wherein the label intrinsic to the target nucleic acid sequence is ^{32}P .

15

58. The method of claim 46 wherein detection of a target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of the heat of hybridization.

20

59. The method of claim 58 wherein a discrete analyte label moiety is linked to each target nucleic acid sequence, and identification of the target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of the discrete analyte label moiety linked to the target nucleic acid sequence.

25

60. The method of claim 45 wherein the sequencing method is by detection of label nucleotide sequences that attach by hybridization to the target nucleic acid sequence.

30

61. The method of claim 60 wherein a discrete analyte label moiety is linked to each target nucleic acid sequence, identification of the label nucleotide sequence hybridizing to a target nucleic acid sequence is by detection of the label moiety, and identification of the target nucleic acid sequence hybridizing to a label nucleotide sequence is by detection of the discrete analyte label moiety linked to the target nucleic acid sequence.

62. The method of claim 61 wherein successive target segments in each target analyte sequence are exposed for hybridization to label nucleotide sequences by endonuclease digestion.

5

63. The method of claim 40 wherein hybridized target nucleic acid sequences are amplified by a polymerase enzyme that requires a hybridized complex for polymerizing the formation of nucleic acid sequence.

10

64. The method of claim 63 wherein hybridized target nucleic acid sequences are amplified by polymerase chain reaction.

65. The method of claim 63 wherein hybridized target nucleic acid sequences are amplified by an RNA replicase enzyme.

15

66. The method of claim 63 employed for genetic analysis.

67. The method of claim 66 employed for allelic analysis.

20

68. The method of claim 66 wherein genomic DNA is analyzed.

69. The method of claim 66 wherein genomic cDNA is analyzed.

70. The method of claim 41 wherein the four nucleotides are A, T, C and G.

25

71. The method of claim 43 wherein the five nucleotides are A, T, C, G and I (Inosine).

72. The method of 45 wherein the sequencing method is by analysis of hybridization data obtained from an array of target nucleic acid sequences attached to a substrate surface.

30

73. The method of claim 72 wherein a discrete analyte label moiety is linked to each target nucleic acid sequence, identification of the label nucleotide sequence hybridizing to a target nucleic acid sequence is by detection of the label moiety, and identification of the target nucleic acid sequence attached to the substrate surface is by
5 detection of the discrete analyte label moiety linked to the target nucleic acid sequence.

74. The method of 72 wherein the array comprises arrayed individual beads or particles, each bead or particle having a surface, the surface having attached a plurality of target nucleic acid sequences having an identical sequence.

10

75. The method of 72 wherein the array comprises arrayed discrete sites on a substrate surface of an integrated substrate, each site having a surface, the surface having attached a plurality of target nucleic acid sequences having an identical sequence.

15

76. The method of claim 72 wherein each label nucleotide sequence additionally comprises a linker moiety.

20

77. The method of claim 76 wherein the linker moiety comprises a common nucleic acid sequence and the label moiety comprises a signature nucleic acid sequence that identifies the target sequence segment.

78. The method of claim 77 wherein the common nucleic acid sequence is double stranded.

25

79. The method of claim 78 wherein decoder labels comprising a nucleic acid sequence complementary to the signature sequence and a second label moiety are employed to image the array.

30

80. The method of claim 79 wherein the second label moiety comprises a luminescent moiety.

81. The method of claim 80 wherein the luminescent moiety is a fluorescent or chemiluminescent moiety.

82. The method of claim 81 wherein the luminescent moiety is phycoerythrin, the double stranded common nucleic acid sequence is 14 nucleotides long the target sequence segment is 4 nucleotides long and the signature sequence is 10 nucleotides long, and successive cycles of hybridization, ligation, detection and endonuclease digestion are employed.

83. The method of claim 72 wherein the substrate surface is functionalized with a surface modification to enhance hybridization.

10

84. The method of claim 83 wherein the enhancement is increasing stringency or kinetics of hybridization.

85. The method of claim 72 wherein the electric potential at the substrate surface is electronically controlled to enhance hybridization.

15

86. The method of claim 75 wherein the integrated substrate comprises a semiconductor chip comprising electronic circuitry, wherein the electric potential at the individual array sites of the substrate surface is independently electronically controlled to enhance hybridization.

20

87. A method for determining a nucleotide at a position of interest in an analyte nucleic acid sequence having a target segment under conditions suitable for hybridization, the method comprising:

25 a) contacting the nucleic acid sequence a first probe and a second probe each probe comprising a nucleic acid sequence complementary, or complementary except at the position of interest, to the target segment,

wherein the position of the first probe corresponding to the position of interest comprises a nucleotide base pairing with a first set of two or more nucleotides present in the nucleic acid sequence and the position of the second probe corresponding to the position of interest comprises a nucleotide base pairing with a second set of two or more nucleotides present in the nucleic acid sequence, and the first set of two or more nucleotides includes at least one nucleotide that is a member of the second set of two or

30

more nucleotides, the two sets are not identical, and a nucleotide present in the nucleic acid sequence of interest is represented in neither set, and further wherein hybridization of the first probe and the second probe to the target segment under the conditions occurs only if complementarity exists between the nucleotide at the position of interest and the nucleotide at the corresponding position of the first probe and the second probe;

b) detecting which of the first probe and the second probe hybridize to the target segment; and

c) analyzing the data from step b),

wherein cumulative information as to the identity of the nucleotide at the position of interest is obtained from the combined data from the first probe and second probe.

88. A method for determining a nucleotide at a position of interest in an nucleic acid sequence having a target segment under conditions suitable for hybridization, the method comprising:

a) contacting the nucleic acid sequence with a first probe comprising a nucleic acid sequence complementary, or complementary except at the position of interest, to the nucleic acid sequence,

wherein the position of the first probe corresponding to the position of interest comprises a nucleotide base pairing with a first set of two or more nucleotides present in the nucleic acid sequence, and that a nucleotide present in the nucleic acid sequence is not represented in the first set of two or more nucleotides, and further wherein hybridization of the first probe to the nucleic acid sequence under the conditions occurs only if complementarity exists between the nucleotide at the position of interest and the nucleotide at the corresponding position of the first probe;

b) contacting the nucleic acid sequence with a second probe comprising a nucleic acid sequence complementary, or complementary except at the position of interest, to the nucleic acid sequence,

provided that the position of the second probe corresponding to the position of interest comprises a nucleotide base pairing with a second set of two or more nucleotides present in the nucleic acid sequence, a nucleotide present in the nucleic acid sequence is not represented in the second set, the second set is different from the first set, and the first set includes at least one nucleotide that is a member of the second set, and the nucleotide that is not represented in the first set is not represented in the second set,

wherein hybridization of the second probe to the nucleic acid sequence under the conditions occurs only if complementarity exists between the nucleotide at the position of interest and the nucleotide at the corresponding position of the second probe;

5 c) detecting whether the first probe and the second probe hybridize to the nucleic acid sequence; and

d) analyzing the hybridization data,

wherein cumulative information as to the identity of the nucleotide at the position of interest is obtained from the combined data from the first and second probes.

10 89. A collection comprised of probe nucleic acid sequence sets, each of the collection of nucleic acid sequence sets having a position corresponding to a probed position of a target nucleic acid sequence,

15 wherein each probed position of nucleic acid sequence set is capable of base pairing to a unique degenerate set of nucleotides, each unique degenerate set of nucleotides has at least one nucleotide in common with each other unique degenerate set of nucleotides, and one nucleotide is commonly excluded from all the unique degenerate sets of nucleotides.

20 90. The collection of claim 89 wherein each sequence set consists of a single sequence.

25 91. The collection of claim 89 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest, a nucleotide base pairing with two nucleotides, and the collection consists of two probe nucleic acid sequence sets.

30 92. The collection of claim 89 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest a nucleotide base pairing with more than two nucleotides, and the collection consists of more than two probe nucleic acid sequences or probe nucleic acid sequence sets.

93. The collection of claim 92 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest a nucleotide base pairing with three nucleotides, and the collection consists of three probe sequence sets.

5 94. The collection of claim 89 in combination with a null probe comprising at the position of interest, complementary to the nucleic acid sequence, having the nucleotide that is commonly excluded from all the unique degenerate sets of nucleotides at the position of interest.

10 95. An array comprising the probe nucleic acid sequences of claim 89 arrayed attached to a substrate surface.

15 96. The array of claim 95 comprising arrayed individual beads or particles, each bead or particle having a surface to which is attached a plurality of probes having an identical sequence.

20 97. The array of claim 95 comprising an integrated substrate having a surface, the surface having a plurality of discrete surface sites, each site having attached a plurality of probe nucleic acid sequences having an identical sequence.

 98. The array of claim 95 wherein each probe nucleic acid sequence additionally comprises a label moiety.

25 99. The collection of claim 89 wherein each probe nucleic acid sequence additionally comprises a label moiety.

 100. The collection of claim 89 wherein each probe nucleic acid sequence additionally comprises a linker moiety and a label moiety.

30 101. The collection of claim 100 wherein the linker moiety comprises a common nucleic acid sequence and the label moiety comprises a signature nucleic acid sequence that identifies the target sequence segment.

102. The collection of claim 101 wherein the common nucleic acid sequence is double stranded.

103. The collection of claim 102 additionally comprising decoders, each decoder
5 comprising a nucleic acid sequence complementary to the signature sequence and a second label moiety.

104. The collection of claim 103 wherein the second label moiety comprises a luminescent moiety.

10

105. The collection of claim 102 wherein the double stranded common nucleic acid sequence is 14 nucleotides long the target segment is 4 nucleotides long and the signature sequence is 10 nucleotides long, and the second label moiety is phycoerythrin.

15

106. The array of claim 95 wherein the substrate surface is functionalized with a surface modification to enhance hybridization.

107. The array of claim 106 wherein the enhancement is increasing stringency or kinetics of hybridization.

20

108. The array of claim 95 wherein the electric potential at the substrate surface is electronically controlled to enhance hybridization.

109. The array of claim 97 wherein the integrated substrate comprises a
25 semiconductor chip comprising electronic circuitry, wherein the electric potential at the individual array sites of the substrate surface is independently electronically controlled to enhance hybridization.

110. A probe system comprising a pair of probe nucleic acid sequence sets, each
30 of the pair of probe nucleic acid sequence sets having a position corresponding to a probed position of a target nucleic acid sequence,

wherein each probed position of each of the pair of probe nucleic acid sequence sets is capable of base pairing to a unique doubly degenerate set of nucleotides, each doubly degenerate set of nucleotides sharing a single common nucleotide.

5 111. The system of claim 110 wherein each sequence set consists of a single sequence.

112. The system of claim 110 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest, a nucleotide base
10 pairing with two nucleotides, and the collection consists of two probe nucleic acid sequence sets.

113. The system of claim 110 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest a nucleotide base
15 pairing with more than two nucleotides, and the collection consists of more than two probe nucleic acid sequences or probe nucleic acid sequence sets.

114. The probe nucleic acid sequences of claim 113 wherein each probe nucleic acid sequence comprises, at the position corresponding to the position of interest a
20 nucleotide base pairing with three nucleotides, and the collection consists of three probe sequence sets.